MAXIMUM-ENTROPY INFERENCE AND INVERSE CONTINUITY OF THE NUMERICAL RANGE

STEPHAN WEIS

ABSTRACT. We study the continuity of the maximum-entropy inference map for two observables in finite dimensions. We prove that the continuity is equivalent to the strong continuity of the set-valued inverse numerical range map. This gives a continuity condition in terms of analytic eigenvalue functions which implies that discontinuities are very rare. It shows also that the continuity of the MaxEnt inference method is independent of the prior state.

Key Words: maximum-entropy inference, continuity, numerical range, strong continuity, stability, strong stability.

2000 Mathematics Subject Classification: Primary 81P16, 62F30, 94A17, 54C10, 47A12, 54C08; Secondary 47N50, 82B26.

1. INTRODUCTION

The maximum-entropy principle, going back to Boltzmann, is one of the standard techniques in quantum mechanical inference problems [16, 48, 15, 39, 2] and state reconstruction [6, 40]. Here we consider a finite set of quantum observables, represented by hermitian matrices in the algebra M_d of complex $d \times d$-matrices, $d \in \mathbb{N}$. If their expected values with respect to several quantum states are identical then no unique quantum state is specified by these expected values. The maximum-entropy inference map makes a definite choice by selecting the state with maximal von Neumann entropy. This inference map from expected values to states can have discontinuity points on the boundary of the set of expected values [47, 45] while analogous inference maps to probability distributions are always continuous. The discontinuities have a meaning in physics. They have been discussed as a signature of a quantum phase transition [8]. They are passed [46, 32] from the inference map to a correlation quantity called irreducible correlation [23, 49] which is connected to the topological entanglement entropy used to characterize topological order [25, 18].

Methods to analyze the discontinuities have included information topology [44], convex geometry [45, 32], and, for two observables, numerical range techniques [32]. Here, we focus on the case of two observables which we encode into a single matrix $A \in M_d$ as its *real part* $\Re(A) := \frac{1}{2}(A+A^*)$ and *imaginary part* $\Im(A) := \frac{1}{2\mathrm{i}}(A-A^*)$, a notation which we will meet again in Sec. 6. The set of *density matrices* in M_d is denoted by

$$\mathcal{M}_d := \{\rho \in M_d \mid \rho \succeq 0, \mathrm{tr}(\rho) = 1\}.$$

This set is also called *state space* [1], $a \succeq 0$ means that the matrix $a \in M_d$ is positive semi-definite. The state space is a convex body, that is a compact convex subset in a Euclidean space. The inner product $\langle a, b\rangle := \mathrm{tr}(a^* b)$, $a, b \in M_d$, and norm $\|a\|_2 := \sqrt{\langle a, a\rangle}$ shall be used.

In quantum mechanics, see for example [4], Secs. 5.1 and 5.2, elements of $\mathcal{M}_d$ represent *states* of a quantum system and the real number $\mathrm{tr}(\rho a) = \langle \rho, a\rangle$, for an observable $a \in M_d$ and for $\rho \in \mathcal{M}_d$, is interpreted as the expected value of a when the system is in the state ρ. Since we use the map $\rho \mapsto \langle \rho, A\rangle$ in various restrictions the notation will simplify by reserving a symbol it. We define the *expected value functional*

$$\mathbb{E}_A : \quad \{b \in M_d \mid b^* = b\} \to \mathbb{C}, \quad a \mapsto \langle a, A\rangle$$

on the Euclidean space of hermitian matrices. The map $\mathbb{E}_A$ sends a state $\rho \in \mathcal{M}_d$ to the pair $\mathbb{E}_A(\rho) = (\langle \rho, \Re(A)\rangle, \langle \rho, \Im(A)\rangle)$ of expected values of the observables $\Re(A)$ and $\Im(A)$, in the identification of the range $\mathbb{C}$ with $\mathbb{R}^2$.

The domain of the maximum-entropy inference map is the convex body

$$L_A := \{\mathbb{E}_A(\rho) \mid \rho \in \mathcal{M}_d\} \subset \mathbb{R}^2,$$

comprising the expected value pairs of $\Re(A)$ and $\Im(A)$. We call L_A *convex support* [47, 45, 32] by its name in probability theory [3]. The *von Neumann entropy* of a state $\rho \in \mathcal{M}_d$ is $S(\rho) = -\mathrm{tr}(\rho \cdot \log \rho)$ and the *maximum-entropy inference* is the map

$$\rho_A^* : \quad L_A \to \mathcal{M}_d, \quad \alpha \mapsto \mathrm{argmax}\{S(\rho) \mid \rho \in \mathcal{M}_d, \mathbb{E}_A(\rho) = \alpha\}.$$

See [16, 15] for more information about ρ_A^*. Our analysis will be based on [45], Thm. 4.9, which affirms that for all $\alpha \in L_A$

$$\text{(1.1)} \quad \rho_A^* \text{ is continuous at } \alpha \text{ if, and only if, } \mathbb{E}_A|_{\mathcal{M}_d} \text{ is open at } \rho_A^*(\alpha).$$

Thereby, a function between topological spaces is *open at* a point in the domain if the image of every neighborhood of that point is a neighborhood of the image point. Clearly, every linear map is open in finite dimensions but it may fail to be open when restricted.

Exact bounds on the number of discontinuity points of ρ^*_A are known for $d \leqslant 5$, see Secs. 7 and 8 in [32]. The bounds have been derived from pre-image results [20] of the following map f_A. The aim of this article is to go beyond these pre-image results and to establish a direct link to a continuity problem in operator theory [10, 21, 24]. Denoting by $S\mathbb{C}^d$ the unit sphere of $\mathbb{C}^d$, the *numerical range map* of a matrix $A \in M_d$ is defined by

$$f_A : \quad S\mathbb{C}^d \to \mathbb{C}, \quad x \mapsto \langle x, Ax\rangle.$$

The *numerical range* is the image $W(A) = f_A(S\mathbb{C}^d)$. Here, $\langle x, y\rangle := \overline{x_1}y_1 + \cdots + \overline{x_d}y_d$, $x, y \in \mathbb{C}^d$, is the inner product of $\mathbb{C}^d$. The numerical range [19, 14] is a convex set by the Toeplitz-Hausdorff theorem [22] and it is well-known that $W(A) = L_A$ holds, see for example [5], Thm. 3. The set-valued inverse $f_A^{-1} : W(A) \to S\mathbb{C}^d$ is called [10, 21, 24] *strongly continuous* at $\alpha \in W(A)$ if for all $x \in f_A^{-1}(\alpha)$ the map f_A is open at x.

Our main result can be summarized as follows.

Theorem 1.1. *For all $\alpha \in L_A$ the maximum-entropy inference map ρ^*_A is continuous at α if and only if f_A^{-1} is strongly continuous at α.*

We remark that there is a very large set of functions which can replace the von Neumann entropy in this continuity analysis. Although the map ρ^*_A will change, its topological properties will remain if $\rho^*_A(\alpha)$ lies in the relative interior of the fiber $\mathbb{E}_A^{-1}(\alpha)$ for all $\alpha \in L_A$ (see Sec. 7 for examples). This is the content of Coro. 5.4 but remains an open problem for more than two observables. Otherwise, if some of the inference points belong to the relative boundary of fibers, the topology can change already for two observables, see [32], Thm. 7.1. We recall that the relative interior of a convex set C is the interior of C in the topology of the affine hull of C. The relative boundary of C is the complement of the relative interior in the closure of C.

The proof of Thm. 1.1 at the end of Sec. 5 uses two properties of convex sets: Firstly (see Sec. 4), like all two-dimensional convex bodies, the set L_A is *stable*, that is the mid-point map $(\alpha, \beta) \mapsto \frac{1}{2}(\alpha + \beta)$ is open [28]. Secondly (see Sec. 5), the state space $\mathcal{M}_d$ is *strongly stable* [35]. These two notions of stability are equivalent for a large class of convex sets [36]. Stability of the set of density matrices on a separable Hilbert space was used, for example, to show the continuity of entanglement monotones [29] arising from the convex roof extension [42]. The present work shows that already in finite dimensions the topology of a linear map on the state space $\mathcal{M}_d$ and the stability of its linear images (see last paragraph of Sec. 2) have much more to explore.

Acknowledgements. Thanks to Lia L. Pugliese for her help proofreading the English version of this essay, to Maksim E. Shirokov for his help with the literature about stability, and to Ilya M. Spitkovsky for his hospitality at the New York University Abu Dhabi (12/2014) and for discussions about eigenvalue functions. Special thanks to Arleta Szkoła and Nihat Ay for general support at the MPI for Mathematics in the Sciences in Leipzig and to the local library there, where I have enjoyed a pleasant working environment (10/2014–05/2015).

2. Remarks and Corollaries

We comment on the main result, derive some corollaries, and provide an outlook.

Thm. 1.1 is surprising because the functions ρ^*_A and f_A^{-1} are opposite in several respects.

Remark 2.1.

(1) Studying the continuity of ρ^*_A requires by (1.1) to check the openness of $\mathbb{E}_A$ restricted to the state space $\mathcal{M}_d$, while studying the strong continuity of f_A^{-1} requires by Lemma 3.2 to check the openness of $\mathbb{E}_A$ restricted to the extremal points $\mathcal{M}_d^1$ of $\mathcal{M}_d$.

(2) Lemma 5.8 in [45] shows that for all $\alpha \in L_A$ the state $\rho^*_A(\alpha)$ lies in the relative interior of the fiber $F := \mathbb{E}_A|_{\mathcal{M}_d}^{-1}(\alpha)$. On the other hand, the set $f_A^{-1}(\alpha)$ consists of extremal points of F. In fact, we have seen in (1) that the elements of $f_A^{-1}(\alpha)$ are extremal points of the state space $\mathcal{M}_d$. *A fortiori* they are extremal points of F.

(3) While for all $\alpha \in L_A$ the state $\rho^*_A(\alpha)$ maximizes the von Neumann entropy on the fiber $\mathbb{E}_A|_{\mathcal{M}_d}^{-1}(\alpha)$, the pre-image $f_A^{-1}(\alpha)$ is the zero level set of minimal entropy, see for example [43], Sec. A.2.

Two corollaries are worth pointing out. Sec. 6 focusses on a continuity condition of f_A^{-1} in terms of analytic eigenvalue functions [21]. The condition then governs the continuity of the inference map ρ^*_A. For example, this shows that ρ^*_A has at most finitely many points of discontinuity and that the set of matrices A where ρ^*_A is continuous is open and dense in M_d.

Sec. 7 addresses the quantum *MaxEnt* inference method [37, 38, 7, 2] which is an updating rule from a prior state to an inference state, given new information in terms of expected values. The maximum-entropy inference is the special case of a uniform prior. We prove that the MaxEnt inference, seen as a function from expected values to inference states, has for all prior states the same points of discontinuity.

Finally, we remark that new ideas will be needed to extend the methods of this article to $r = 3$ (or more) observables $u_1, \ldots, u_r \in M_d$. Firstly, it was observed in [8], Exa. 6 (see also [32], Exa. 5.2) that the convex support $\{(\langle\rho, u_1\rangle, \ldots, \langle\rho, u_r\rangle) \mid \rho \in \mathcal{M}_d\}$ is not stable for some choices of observables $u_1, u_2, u_3 \in M_3$ when $r = d = 3$. Secondly, although the joint numerical range $\{\langle x, u_i x\rangle_{i=1}^r \mid x \in S\mathbb{C}^d\}$ contains the extremal points of the convex support, it is in general not convex [22] for $r \geqslant 3$. So the boundary of the convex support will need a careful analysis when trying to go beyond $r = 2$.

3. Preliminaries

We introduce faces of convex sets and pure states. We connect the domains of the functions f_A and $\mathbb{E}_A|_{\mathcal{M}_d}$ by recalling properties of the quotient map $\beta : S\mathbb{C}^d \to \mathbb{P}\mathbb{C}^d$ from the unit sphere in $\mathbb{C}^d$ to the projective space of lines in $\mathbb{C}^d$. This is a well-known smooth (even real analytic) map. Nevertheless we provide a proof because we are also interested in the openness of β.

A *face* of a convex set C is a convex subset $F \subset C$ such that if for $x, y \in C$ the open segment $]x, y[:= \{(1-\lambda)x + \lambda y \mid \lambda \in (0,1)\}$ intersects F, then the closed segment $[x, y] := \{(1-\lambda)x + \lambda y \mid \lambda \in [0,1]\}$ belongs to F. An *extremal point* is a face of dimension zero and a *facet* is a face of dimension $\dim(C) - 1$.

The extremal points of the state space $\mathcal{M}_d$, $d \in \mathbb{N}$, are called *pure states* in physics [4, 27] and it is well-known, see for example (4.2) in [1], that the set of pure states equals the set of rank-one density matrices which we denote by

$$\mathcal{M}_d^1 := \{\rho \in \mathcal{M}_d \mid \mathrm{rank}(\rho) = 1\}. \tag{3.1}$$

The rank-one density matrices are the orthogonal projections onto one-dimensional subspaces of $\mathbb{C}^d$. So $\mathcal{M}_d^1 \cong \mathbb{P}\mathbb{C}^d = S\mathbb{C}^d / S\mathbb{C}^1$ is a projective space. We denote the quotient map in Dirac's notation

$$\beta : \quad S\mathbb{C}^d \to \mathcal{M}_d^1, \quad x \mapsto |x\rangle\langle x|. \tag{3.2}$$

Its fibers are isomorphic to the circle $S\mathbb{C}^1$. For $d = 2$ the famous Hopf fibration is obtained, see for example [4].

In the following we use the trace distance and the fidelity, see for example [27], Sec. 9.2.1–2. Let $\sqrt{a}$ denote the square root of a positive semi-definite matrix $a \in M_d$, that is $\sqrt{a} \succeq 0$ and $(\sqrt{a})^2 = a$. The *trace norm* of $a \in M_d$ is $\|a\|_1 := \mathrm{tr}\sqrt{a^* a}$. The *trace distance* between states $\rho, \sigma \in \mathcal{M}_d$ is

$$D(\rho, \sigma) := \tfrac{1}{2}\|\rho - \sigma\|_1$$

and their *fidelity* is $F(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1 = \mathrm{tr}\sqrt{\sqrt{\rho}\,\sigma\sqrt{\rho}}$.

The fidelity is symmetric in the two arguments by Uhlmann's theorem [41]. We have $0 \leqslant F(\rho,\sigma) \leqslant 1$ where the upper bound is achieved if and only if $\rho = \sigma$. The Fuchs-van de Graaf inequalities [11]

$$1 - F(\rho,\sigma) \leqslant D(\rho,\sigma) \leqslant \sqrt{1 - F(\rho,\sigma)^2} \tag{3.3}$$

hold. For pure states we have $F(|x\rangle\langle x|, |y\rangle\langle y|) = |\langle x, y\rangle|$, $x, y \in S\mathbb{C}^d$, where $|z| := \sqrt{\langle z, z\rangle}$ is the norm of $z \in \mathbb{C}^d$.

We say a function between topological spaces is *open on* a subset of its domain if it is open at each point of this subset. The function is *open* if it is open on the domain.

Lemma 3.1. *The map $\beta : S\mathbb{C}^d \to \mathcal{M}_d^1$ is continuous and open.*

Proof: The second inequality in (3.3) shows $D^2 \leqslant (1-F^2) \leqslant 2(1-F)$. So, for $x, y \in S\mathbb{C}^d$ we have

$$\begin{aligned} D(|x\rangle\langle x|, |y\rangle\langle y|)^2 &\leqslant |x|^2 + |y|^2 - 2|\langle x, y\rangle| \leqslant |x|^2 + |y|^2 - 2\Re(\langle x, y\rangle) \\ &= |x - y|^2, \end{aligned}$$

whence β is Lipschitz-continuous with the global constant one. The left-hand side inequality in (3.3) implies for all $x, y \in S\mathbb{C}^d$ such that $\langle x, y\rangle \geqslant 0$ the inequality of

$$|x - y|^2 = 2(1 - |\langle x, y\rangle|) = 2(1 - F(|x\rangle\langle x|, |y\rangle\langle y|)) \leqslant 2D(|x\rangle\langle x|, |y\rangle\langle y|).$$

This proves that the ball in $S\mathbb{C}^d$ of (Hilbert space) radius $\epsilon > 0$ about $x \in S\mathbb{C}^d$, mapped through β, contains the ball in $\mathcal{M}_d^1$ of (trace distance) radius $\frac{1}{2}\epsilon^2$ about $|x\rangle\langle x|$. Hence β is open. □

Turning to the convex support and to the numerical range we notice that for all $x \in S\mathbb{C}^d$

$$f_A(x) = \langle x, Ax\rangle = \mathrm{tr}(|x\rangle\langle x|A) = \mathbb{E}_A(|x\rangle\langle x|) = \mathbb{E}_A \circ \beta(x). \tag{3.4}$$

Lemma 3.2. *For all $x \in S\mathbb{C}^d$ the following statements are equivalent.*

(1) *The map f_A is open at x.*

(2) *The map $\mathbb{E}_A|_{\mathcal{M}_d^1}$ is open at $|x\rangle\langle x|$.*

Proof: Using (3.4), (1) $\Longrightarrow$ (2) follows from the continuity of β and (2) $\Longrightarrow$ (1) follows from the openness, proved in Lemma 3.1. □

4. Strong Continuity Implies Continuity

We prove that the strong continuity of f_A^{-1} implies the continuity of the maximum entropy inference ρ_A^*. The main argument is the stability of two-dimensional convex bodies.

A convex body C is *stable* if $C \times C \to C$, $(x,y) \mapsto \frac{1}{2}(x+y)$ is an open map [28, 9]. We recall two facts about stable convex bodies. Firstly, if C is a stable convex body then for any integer $n \geqslant 2$ and for $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ such that $\lambda_i \geqslant 0$ for $i = 1, \dots, n$ and $\lambda_1 + \dots + \lambda_n = 1$ the map

$$\underbrace{C \times \cdots \times C}_{n \text{ times}} \to C, \quad (x_1, \dots, x_n) \mapsto \lambda_1 x_1 + \cdots + \lambda_n x_n \tag{4.1}$$

is open. The proof that (4.1) is open is given for $n = 2$ in [9], Prop. 1.1, and the case of $n \geqslant 3$ follows by induction.

Secondly, every two-dimensional convex body is stable. This follows from Thm. 2.3 in [28] which says that a convex body C of any finite dimension $l \in \mathbb{N}$ is stable if, and only if, for each $k = 0, \dots, l$ the k-*skeleton*, that is the union of all faces of C of dimension at most k, is closed. The $(l-2)$-, the $(l-1)$- and the l-skeletons of C are always closed, see [12], so every two-dimensional convex body is stable.

The Euclidean ball of radius $\epsilon > 0$ about $\rho \in \mathcal{M}_d$ within a subset $C \subset \mathcal{M}_d$ will be denoted by $B_\epsilon(\rho, C) := \{\sigma \in C \mid \|\rho - \sigma\|_2 \leqslant \epsilon\}$.

Theorem 4.1. *Let α be an extremal point of L_A. If f_A^{-1} is strongly continuous at α then $\mathbb{E}_A|_{\mathcal{M}_d}$ is open on $\mathbb{E}_A|_{\mathcal{M}_d}^{-1}(\alpha)$.*

Proof: If $\alpha \in L_A$ is an extremal point then the fiber $F := \mathbb{E}_A|_{\mathcal{M}_d}^{-1}(\alpha)$ is a face of the state space $\mathcal{M}_d$, so all extremal points of F are pure states or equivalently, by (3.1), they belong to the set of rank-one states $\mathcal{M}_d^1$. Hence we can write an arbitrary point $\rho \in F$ in the form

$$\rho = \lambda_1 \rho_1 + \cdots + \lambda_n \rho_n$$

where $\rho_i \in \mathcal{M}_d^1 \cap F$, $\lambda_i \geqslant 0$ for $i = 1, \dots, n$ and $\lambda_1 + \cdots + \lambda_n = 1$. Let $x_i \in S\mathbb{C}^d$ such that $\rho_i = |x_i\rangle\langle x_i|$ and choose a neighborhood $N_i \subset \mathcal{M}_d^1$ of ρ_i in $\mathcal{M}_d^1$. By the continuity of $\beta : S\mathbb{C}^d \to \mathcal{M}_d^1$ (see Lemma 3.1) the pre-image $N_i' := \beta^{-1}(N_i)$ is a neighborhood of x_i in $S\mathbb{C}^d$. The assumption that f_A^{-1} is strongly continuous at α proves that $f_A(N_i')$ is a neighborhood of α. Hence $\mathbb{E}_A(N_i) = f_A(N_i')$ is a neighborhood of α.

Now let $N \subset \mathcal{M}_d$ be an arbitrary neighborhood of ρ in $\mathcal{M}_d$ and choose neighborhoods $N_i \subset \mathcal{M}_d^1$ about ρ_i in $\mathcal{M}_d^1$ such that $\lambda_1 N_1 + \cdots + \lambda_n N_n \subset N$. It suffices to consider a Euclidean ball $B_\epsilon(\rho, \mathcal{M}_d) \subset N$ of radius $\epsilon > 0$ about ρ and to use the Euclidean balls $N_i = B_\epsilon(\rho_i, \mathcal{M}_d^1)$

about ρ_i, $i = 1, \dots, n$. Then

$$\lambda_1 \mathbb{E}_A(N_1) + \cdots + \lambda_n \mathbb{E}_A(N_n) \subset \mathbb{E}_A(N).$$

We have seen that each set $\mathbb{E}_A(N_i)$ is a neighborhood of α and we have pointed out earlier in this section that the two-dimensional convex body L_A is stable. Hence (4.1) shows that $\mathbb{E}_A(N)$ contains a neighborhood of α. This completes the proof. □

The linear map $\mathbb{E}_A|_{\mathcal{M}_d}$ is open on the fiber $\mathbb{E}_A|_{\mathcal{M}_d}^{-1}(\alpha)$ of $\alpha \in L_A$ if α is a relative interior point of L_A or a relative interior point of a facet of L_A. For a proof see [45], Sec. 4.3, or [32], Sec. 3. Since $\dim(L_A) \leqslant 2$ we deduce from Thm. 4.1 the following.

Corollary 4.2. *If f_A^{-1} is strongly continuous at $\alpha \in L_A$ then $\mathbb{E}_A|_{\mathcal{M}_d}$ is open on the fiber $\mathbb{E}_A|_{\mathcal{M}_d}^{-1}(\alpha)$.*

5. Continuity Implies Strong Continuity

We prove that the continuity of ρ_A^* implies the strong continuity of f_A^{-1}. This is the harder part compared to converse direction in Sec. 4 because we now have to restrict the domain from the state space $\mathcal{M}_d$ to the pure states $\mathcal{M}_d^1$ while keeping the range L_A. A major argument will be a corollary of the *strong stability* of the state space [35].

It is well-known that the state space $\mathcal{M}_d$ is stable. Indeed, Lemma 3 in [34] proves that the map

$$\mathcal{M}_d \times \mathcal{M}_d \times [0,1] \to \mathcal{M}_d, \quad (\rho, \sigma, \lambda) \mapsto (1-\lambda)\rho + \lambda\sigma \tag{5.1}$$

is open, which is equivalent to the stability of $\mathcal{M}_d$ by Prop. 1.1 in [9]. To make an openness statement about $\mathbb{E}_A|_{\mathcal{M}_d^1}$ we have to restrict the left-hand side of (5.1) from $\mathcal{M}_d$ to $\mathcal{M}_d^1$ while keeping the right-hand side. This restriction is indeed possible. The cost is the non-finiteness of the ensemble, see Rem. 1 in [35]. The corresponding property of $\mathcal{M}_d$ is called *strong stability* which, by definition, means that for all $k = 1, \dots, d$ the barycenter map from the discrete probability measures on $\{\rho \in \mathcal{M}_d \mid \operatorname{rank}(\rho) \leqslant k\}$ to $\mathcal{M}_d$ is open, see [35], Thm. 1.

Lemma 4 in [35] serves for our purposes: Let $\{\pi_i, \rho_i\}_{i\in\mathbb{N}}$ be a countable ensemble, that is $\rho_i \in \mathcal{M}_d^1$, $\pi_i \geqslant 0$ for all $i \in \mathbb{N}$ and $\sum_{i=1}^\infty \pi_i = 1$. For an arbitrary sequence $\{\rho^n\} \subset \mathcal{M}_d$ converging to the average $\sum_{i=1}^\infty \pi_i\rho_i$ there exists a sequence $\{\{\pi_i^n, \rho_i^n\}_{i\in\mathbb{N}}\}_{n\in\mathbb{N}}$ of countable ensembles such that

$$\begin{array}{ll} (\forall n) & \pi_1^n\rho_1^n + \pi_2^n\rho_2^n + \cdots = \rho^n, \\ (\forall i) & \lim_{n\to\infty} \pi_i^n = \pi_i \quad \text{and} \quad (\pi_i > 0 \implies \lim_{n\to\infty} \rho_i^n = \rho_i). \end{array} \tag{5.2}$$

We use an immediate corollary of (5.2) which is as follows.

Corollary 5.1. *Let $\rho \in \mathcal{M}_d^1$ be a pure state, let $N \subset \mathcal{M}_d^1$ be a neighborhood of ρ in $\mathcal{M}_d^1$ and let $\sigma \in \mathcal{M}_d$ be an arbitrary state. For every $\lambda \in [0,1)$ and every $\tilde{\lambda} > \lambda$ with $\tilde{\lambda} \leqslant 1$ the set $(1-\tilde{\lambda})N + \tilde{\lambda}\mathcal{M}_d$ is a neighborhood of $(1-\lambda)\rho + \lambda\sigma$ in $\mathcal{M}_d$.*

The Bloch ball $\mathcal{M}_2$ with $\rho = |0\rangle\langle 0|$, $\sigma = |1\rangle\langle 1|$, and $\lambda \leqslant \frac{1}{2}$ shows that the assumption $\tilde{\lambda} > \lambda$ of Coro. 5.1 can not be weakened to $\tilde{\lambda} \geqslant \lambda$.

Based on two chapters of the numerical range theory, the next lemma provides an extremal point argument. Firstly, Thm. 2 in [10] states that for all $\alpha \in L_A$ and $x \in S\mathbb{C}^d$ such that $\alpha = f_A(x)$ and for any neighborhood U of x in $S\mathbb{C}^d$ there is a constant $\delta > 0$ such that $\delta L_A + (1-\delta)\alpha \subset f_A(U)$ holds. Secondly, Lemma 3.2 in [21] proves that for $r > 0$ and $x \in S\mathbb{C}^d$ the set $f_A(\{y \in S\mathbb{C}^d \mid |y - x| < r\})$ is convex.

Lemma 5.2. *Let $\rho \in \mathcal{M}_d^1$ and let $N \subset \mathcal{M}_d^1$ be a neighborhood of ρ in $\mathcal{M}_d^1$. There exists a neighborhood $\widetilde{N} \subset N$ of ρ in $\mathcal{M}_d^1$ such that $\mathbb{E}_A(\widetilde{N})$ is convex. The set $\mathbb{E}_A(\widetilde{N})$ is a neighborhood of $\mathbb{E}_A(\rho)$ in L_A if it contains all extremal points of L_A in a neighborhood of $\mathbb{E}_A(\rho)$ in L_A.*

Proof: The continuity of the quotient map $\beta : S\mathbb{C}^d \to \mathcal{M}_d^1$, see Lemma 3.1, shows that $N' := \beta^{-1}(N) \subset S\mathbb{C}^d$ is a neighborhood of any point in $\beta^{-1}(\rho)$. Let x be such a point. Lemma 3.2 in [21], cited above, shows that for some neighborhood $N'' \subset N'$ of x the image $f_A(N'')$ is convex. The openness of β shows that $\widetilde{N} := \beta(N'')$ is a neighborhood of ρ in $\mathcal{M}_d^1$. Moreover, $\mathbb{E}_A(\widetilde{N}) = f_A(N'')$ is convex and contains $\alpha := \mathbb{E}_A(\rho) = f_A(x)$. This proves the first assertion.

Let us prove that $f_A(N'')$ is a neighborhood of α if it contains all extremal points of L_A sufficiently close to α. We can assume that L_A has real dimension two and that α is an extremal point of L_A. Otherwise Thm. 2 in [10], cited above, shows that $f_A(N'')$ is a neighborhood of α.

Since $f_A(N'')$ is convex, it suffices to show that it contains a neighborhood of α in ∂L_A. We consider a disk $D := \{\alpha' \in \mathbb{C} \mid |\alpha' - \alpha| < \epsilon\}$ of radius $\epsilon > 0$ about α and one of the semi-arcs, denoted by r_ϵ, on the curve $D \cap \partial L_A$ and starting at α. If r_ϵ is a segment for some ϵ then by Thm. 2 in [10] $f_A(N'')$ contains a neighborhood of α in r_ϵ. Otherwise r_ϵ has a sequence of extremal points of L_A which converges to α. By assumptions, $f_A(N'')$ includes all extremal points of L_A which are sufficiently close to α. Hence, the convex set $f_A(N'')$ contains the segments between those extremal points and therefore a neighborhood of α in r_ϵ. Together with the analogous statement about the other semi-arc we have shown that $f_A(N'')$ contains a neighborhood of α. □

We are ready to prove a main result of this paper.

Theorem 5.3. *Let $\alpha \in L_A$. If $\mathbb{E}_A|_{\mathcal{M}_d}$ is open at a relative interior point of $\mathbb{E}_A|_{\mathcal{M}_d}^{-1}(\alpha)$ then f_A^{-1} is strongly continuous at α.*

Proof: Let ρ be a relative interior point of the fiber $F := \mathbb{E}_A|_{\mathcal{M}_d}^{-1}(\alpha)$ such that $\mathbb{E}_A|_{\mathcal{M}_d}$ is open at ρ. We have to prove that f_A is open at every point of $f_A^{-1}(\alpha)$. Thus, using Lemma 3.2, it suffices to prove that for all pure states $\sigma \in F$ the map $\mathbb{E}_A|_{\mathcal{M}_d^1}$ is open at σ.

Let $N \subset \mathcal{M}_d^1$ be a neighborhood of σ in the set of pure states $\mathcal{M}_d^1$. By Lemma 5.2 there exists a neighborhood $N' \subset N$ of σ in $\mathcal{M}_d^1$ such that $\mathbb{E}_A(N')$ is a neighborhood of α in L_A provided that it contains all extremal points of L_A close to α, which we shall prove now.

Let $\tau \in F$ and $\lambda \in (0,1)$ such that $\rho = (1-\lambda)\sigma + \lambda\tau$, and let $\tilde{\lambda} > \lambda$ with $\tilde{\lambda} < 1$. Then Coro. 5.1 proves that

$$N'' := (1-\tilde{\lambda})N' + \tilde{\lambda}\mathcal{M}_d$$

is a neighborhood of ρ in $\mathcal{M}_d$. By assumptions, $\mathbb{E}_A|_{\mathcal{M}_d}$ is open at ρ, so

$$\mathbb{E}_A(N'') = (1-\tilde{\lambda})\mathbb{E}_A(N') + \tilde{\lambda}L_A \tag{5.3}$$

is a neighborhood of α in L_A. The definition of extremal points and (5.3) show that every extremal point of L_A which lies in $\mathbb{E}_A(N'')$ must lie in $\mathbb{E}_A(N')$. □

Thm. 5.3 and Coro. 4.2 prove the following.

Corollary 5.4. *Let $\alpha \in L_A$. Then $\mathbb{E}_A|_{\mathcal{M}_d}$ is open at a relative interior point of $\mathbb{E}_A|_{\mathcal{M}_d}^{-1}(\alpha)$ if and only if f_A^{-1} is strongly continuous at α. In this case $\mathbb{E}_A|_{\mathcal{M}_d}$ is open on $\mathbb{E}_A|_{\mathcal{M}_d}^{-1}(\alpha)$.*

We are ready to prove Theorem 1.1.

Proof of Theorem 1.1: As we have recalled in Rem. 2.1(2), the inference state $\rho_A^*(\alpha)$ lies in the relative interior of $\mathbb{E}_A|_{\mathcal{M}_d}^{-1}(\alpha)$ for all $\alpha \in L_A$. Therefore the claim follows from (1.1) and Coro. 5.4. □

6. Continuity in Terms of Eigenvalue Functions

We derive a continuity condition of the maximum-entropy inference from the theory of the numerical range map [21]. This shows that discontinuities of the maximum-entropy inference are the exception.

For all $\theta \in \mathbb{R}$ the hermitian matrix

$$\Re(e^{-\mathrm{i}\theta}A) = \cos(\theta)\Re(A) + \sin(\theta)\Im(A)$$

has an orthogonal basis of eigenvectors $\{x_k(\theta)\}_{k=1}^d$ and eigenvalues $\{\lambda_k(\theta)\}_{k=1}^d$ which depend real analytically on θ [30]. Further, we define curves for $k = 1, \ldots, d$,

$$z_k(\theta) := e^{\mathrm{i}\,\theta}(\lambda_k(\theta) + \mathrm{i}\,\lambda'_k(\theta)), \qquad \theta \in \mathbb{R}, \tag{6.1}$$

where λ'_k is the derivative of λ_k with respect to θ. We remark that the union of these curves is a plane algebraic curve [17] whose convex hull is the numerical range [19].

An eigenvalue function λ_k is said [21] to *correspond* to $\alpha \in W(A)$ at $\theta \in \mathbb{R}$ if $z_k(\theta) = \alpha$. Notice by (6.1), all eigenvalue functions corresponding to $\alpha \in W(A)$ at $\theta \in \mathbb{R}$ have the same value and the same derivative at θ. We denote the support line of $W(A)$ with outward pointing normal vector $-e^{\mathrm{i}\,\theta}$ by ℓ_θ. The following Fact 6.1(1) is proved in Thm. 2.1(1) in [21]. Part (2) follows from Thm. 2 in [10].

Fact 6.1. (1) *Let $\theta \in \mathbb{R}$ and let $\alpha \in W(A) \cap \ell_\theta$ be an extremal point of $W(A)$. Then f_A^{-1} is strongly continuous at α if and only if the eigenvalue functions corresponding to α at θ are mutually equal.*

(2) *Condition* (1) *is for all points of $W(A)$ decisive, because f_A^{-1} is strongly continuous at relative interior points of $W(A)$ and at relative interior points of facets.*

An example and two corollaries will illustrate the use of Fact 6.1.

Example 6.2. A discontinuity of the maximum-entropy inference ρ_A^* is known [47, 8, 32] for

$$A := \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \oplus [\,1\,],$$

the direct sum denoting a block-diagonal matrix in M_3. The numerical range $W(A)$ is the unit disk in $\mathbb{C}$ where ρ_A^* is discontinuous at 1.

Let us derive this discontinuity with new methods. The real part of the matrix $e^{-\mathrm{i}\,\theta}A$ is

$$\Re(e^{-\mathrm{i}\,\theta}A) = (\cos(\theta)\sigma_1 + \sin(\theta)\sigma_2) \oplus \cos(\theta), \qquad \theta \in \mathbb{R},$$

for Pauli matrices

$$\sigma_1 := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad \sigma_2 := \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix}.$$

The eigenvalue functions are $\lambda_1(\theta) = 1$, $\lambda_2(\theta) = -1$ and $\lambda_3(\theta) = \cos(\theta)$ while $z_1(\theta) = e^{\mathrm{i}\,\theta}$, $z_2(\theta) = -e^{\mathrm{i}\,\theta}$ and $z_3(\theta) = 1$ holds. The eigenvalue functions corresponding to 1 at π are λ_2 and λ_3 and $1 \in \ell_\pi$ holds. Since $\lambda_2 \neq \lambda_3$, Fact 6.1 proves that f_A^{-1} is not strongly continuous at 1. Thus, Theorem 1.1 proves the discontinuity of ρ_A^* at 1.

Finally we deduce, following [21], that a discontinuity of the maximum-entropy inference is the exception. The eigenvalue functions λ_k extend analytically to a neighborhood of $\mathbb{R}$ in $\mathbb{C}$ and therefore, see for example [33], Thm. 10.18, distinct eigenvalue functions can only coincide at finitely many exceptional values of $\theta \in [0, 2\pi)$. Thus Fact 6.1, Coro. 5.4 and Thm. 1.1 show the following.

Corollary 6.3.

(1) *For all except possibly finitely many points α of L_A the map $\mathbb{E}_A|_{\mathcal{M}_d}$ is open on $\mathbb{E}_A|_{\mathcal{M}_d}^{-1}(\alpha)$.*

(2) *For all except possibly finitely many points α of L_A the maximum-entropy inference ρ_A^* is continuous at α.*

To make a statement about exceptionality of discontinuities in terms of observables we refrain from the global assumption that a matrix $A \in M_d$ is chosen. For all $d \in \mathbb{N}$ the von Neumann-Wigner non-crossing rule [26] in the formulation of [13], Prop. 4.9, states that the set of matrices $A \in M_d$ such that for any $(s,t) \in \mathbb{R}^2 \backslash \{0\}$ the hermitian matrix $s\Re(A) + t\Im(A)$ has simple eigenvalues is open and dense in M_d. Thus Fact 6.1, Coro. 5.4 and Thm. 1.1 show the following.

Corollary 6.4. *Let $d \in \mathbb{N}$.*

(1) *The set of matrices $A \in M_d$ where $\mathbb{E}_A|_{\mathcal{M}_d}$ is open is open and dense in M_d.*

(2) *The set of matrices $A \in M_d$ where ρ_A^* is continuous is open and dense in M_d.*

7. Independence of the Prior State

The *MaxEnt* inference method [37, 38, 7, 2] is an updating rule from a prior state to an inference state, given new information in terms of expected values. We show that the continuity of the MaxEnt inference function does not depend on the prior state.

As we have seen in Sec. 2 the set of expected values of two observables is the convex support L_A which refers to a matrix $A \in M_d$, $d \in \mathbb{N}$. By definition, the prior state $\rho \in \mathcal{M}_d$ is assumed to be a positive definite matrix. The *MaxEnt* inference function, with respect to A and ρ, is

$$\Psi_{A,\rho}: \quad L_A \to \mathcal{M}_d, \quad \alpha \mapsto \operatorname{argmin}\{S(\sigma,\rho) \mid \sigma \in \mathbb{E}_A|_{\mathcal{M}_d}^{-1}(\alpha)\}.$$

This is a well-defined single-valued function [45]. Here, the *Umegaki relative entropy* $S : \mathcal{M}_d \times \mathcal{M}_d \to [0,\infty]$ is an asymmetric distance which is zero only for equal arguments. By definition, for states $\rho \in \mathcal{M}_d$ of maximal rank holds $S(\sigma,\rho) = \operatorname{tr}\sigma(\log(\sigma) - \log(\rho))$. Notice $S(\sigma, \mathbb{1}/d) =$

$\log(d) - S(\sigma)$ for all $\sigma \in \mathcal{M}_d$ where $S(\sigma)$ is the von Neumann entropy. So, $\Psi_{A,\mathbb{1}/d} = \rho_A^*$ is the maximum-entropy inference. Here $\mathbb{1}$ denotes the $d \times d$ identity matrix.

The question whether the continuity of $\Psi_{A,\rho}$ depends on ρ was asked in [45], Rem. 5.9. For two observables the answer is negative:

Theorem 7.1. *All maps in the set* $\{\Psi_{A,\rho} \mid \rho \text{ is a prior state }\}$ *have the same points of discontinuity in* L_A.

Proof: Since the function $\sigma \mapsto S(\sigma, \rho)$ is continuous for each positive definite prior state $\rho \in \mathcal{M}_d$, the continuity of $\Psi_{A,\rho}$ at $\alpha \in L_A$ is equivalent to the openness of $\mathbb{E}_A|_{\mathcal{M}_d}$ at $\Psi_{A,\rho}(\alpha)$, see [45], Thm. 4.9. In addition, Lemma 5.8 in [45] proves that $\Psi_{A,\rho}(\alpha)$ lies in the relative interior of the fiber $\mathbb{E}_A|_{\mathcal{M}_d}^{-1}(\alpha)$. Now Coro. 5.4 proves the claim. □

References

[1] E. M. Alfsen, F. W. Shultz (2001) *State Spaces of Operator Algebras: Basic Theory, Orientations, and C*-Products*, Springer-Verlag

[2] S. A. Ali, C. Cafaro, A. Giffin, C. Lupo, S. Mancini (2012) *On a differential geometric viewpoint of Jaynes' MaxEnt method and its quantum extension*, AIP Conf. Proc. 1443 120-128

[3] O. Barndorff-Nielsen (1978) *Information and Exponential Families in Statistical Theory*, John Wiley & Sons, New York

[4] I. Bengtsson, K. Życzkowski (2006) *Geometry of Quantum States*, Cambridge University Press

[5] S. K. Berberian, G. H. Orland (1967) *On the closure of the numerical range of an operator*, Proc Amer Math Soc 18(3) 499-503

[6] V. Bužek, G. Drobný, R. Derka, G. Adam, H. Wiedemann (1999) *Quantum state reconstruction from incomplete data*, Chaos, Solitons & Fractals 10 981–1074

[7] A. Caticha, A. Giffin (2006) *Updating Probabilities*, AIP Conf. Proc. 872 31–42

[8] J. Chen, Z. Ji, C.-K. Li, Y.-T. Poon, Y. Shen, N. Yu, B. Zeng, D. Zhou (2015) *Discontinuity of maximum entropy inference and quantum phase transitions*, `arXiv:1406.5046v2 [quant-ph]`

[9] A. Clausing, S. Papadopoulou (1978) *Stable convex sets and extremal operators*, Mathematische Annalen 231 193–203

[10] D. Corey, C. R. Johnson, R. Kirk, B. Lins, I. Spitkovsky (2013) *Continuity properties of vectors realizing points in the classical field of values*, Linear and Multilinear Algebra 61(10) 1329–1338

[11] C. A. Fuchs, J. Van De Graaf (1999) *Cryptographic distinguishability measures for quantum-mechanical states*, IEEE Trans Inf Theory 45(4) 1216–1227

[12] R. Grzaslewicz (1997) *Extreme continuous function property*, Acta Mathematica Hungarica 74(1–2) 93–99

[13] E. Gutkin, E. A. Jonckheere, M. Karow (2004) *Convexity of the joint numerical range: topological and differential geometric viewpoints*, Linear Algebra and its Applications 376 143–171

[14] R. A. Horn, C. R. Johnson (1991) *Topics in Matrix Analysis*, Cambridge University Press, Cambridge
[15] R. S. Ingarden, A. Kossakowski, M. Ohya (1997) *Information Dynamics and Open Systems*, Kluwer Academic Publishers Group
[16] E. T. Jaynes (1957) *Information theory and statistical mechanics.*, Phys Rev 106 620–630 and 108 171–190
[17] M. Joswig, B. Straub (1998) *On the numerical range map*, Journal of the Australian Mathematical Society 65 267–283
[18] K. Kato, F. Furrer, M. Murao (2015) *Equivalence of topological entanglement entropy and irreducible correlation and relationship to secret sharing*, `arXiv:1505.01917 [quant-ph]`
[19] R. Kippenhahn (1951) *Über den Wertevorrat einer Matrix*, Mathematische Nachrichten 6(3–4) 193–228
[20] T. Leake, B. Lins, I. M. Spitkovsky (2014) *Pre-images of boundary points of the numerical range*, Operators and Matrices 8(3) 699–724
[21] T. Leake, B. Lins, I. M. Spitkovsky (2014) *Inverse continuity on the boundary of the numerical range*, Linear and Multilinear Algebra 62 1335–1345
[22] C.-K. Li, Y.-T. Poon (2000) *Convexity of the joint numerical range*, SIAM J Matrix Anal A 21(2) 668–678
[23] N. Linden, S. Popescu, W. Wootters (2002) *Almost every pure state of three qubits is completely determined by its two-particle reduced density matrices*, Physical Review Letters 89(20) 207901
[24] B. Lins, P. Parihar (2015) *Continuous selections of the inverse numerical range map*, Linear and Multilinear Algebra 1–13
[25] Y. Liu, B. Zeng, D. L. Zhou (2014) *Irreducible many-body correlations in topologically ordered systems*, `arXiv:1402.4245 [quant-ph]`
[26] J. von Neumann, E. P. Wigner (1929) *Über das Verhalten von Eigenwerten bei adiabatischen Prozessen*, Physikalische Zeitschrift 30 467–470
[27] M. A. Nielsen, I. L. Chuang (2000) *Quantum Computation and Quantum Information*, Cambridge University Press
[28] S. Papadopoulou (1977) *On the geometry of stable compact convex sets*, Math Ann 229 193–200
[29] V. Y. Protasov, M. E. Shirokov (2009) *Generalized compactness in linear spaces and its applications*, Sbornik: Mathematics 200(5) 697–722
[30] F. Rellich (1954) *Perturbation Theory of Eigenvalue Problems*, Research in the Field of Perturbation Theory and Linear Operators, Technical Report No. 1, Courant Institute of Mathematical Sciences, New York University
[31] R. T. Rockafellar (1972) *Convex Analysis*, Princeton University Press
[32] L. Rodman, I. M. Spitkovsky, A. Szkoła, S. Weis (2016) *Continuity of the maximum-entropy inference: Convex geometry and numerical ranges approach*, Journal of Mathematical Physics 57, 015204
[33] W. Rudin (1987) *Real and Complex Analysis*, 3rd ed., McGraw-Hill
[34] M. E. Shirokov (2006) *The Holevo capacity of infinite dimensional channels and the additivity problem*, Commun Math Phys 262 137–159
[35] M. E. Shirokov (2010) *Continuity of the von Neumann entropy*, Commun Math Phys 296(3) 625–654
[36] M. E. Shirokov (2012) *Stability of convex sets and applications*, Izvestiya: Mathematics 76(4) 840–856

[37] J. E. Shore, R. W. Johnson (1980) *Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy*, IEEE Trans. Inf. Theory 26 26–37; Correction (1983) ibid. 29, 942–943

[38] J. Skilling (1989) *Classic maximum entropy*, Maximum Entropy and Bayesian Methods, Springer 45–52.

[39] R. F. Streater (2011) *Proof of a modified Jaynes's estimation theory*, Open Systems & Information Dynamics 18(2) 223–233

[40] B. Swingle, I. H. Kim (2014) *Reconstructing quantum states from local data*, Physical Review Letters 113(26) 2014

[41] A. Uhlmann (1976) *The "transition probability" in the state space of a *-algebra*, Reports on Mathematical Physics 9(2) 273–279

[42] A. Uhlmann (2010) *Roofs and convexity*, Entropy 12(7) 1799–1832

[43] A. Wehrl (1978) *General properties of entropy*, Reviews of Modern Physics 50 221–260

[44] S. Weis (2014) *Information topologies on non-commutative state spaces*, Journal of Convex Analysis 21(2) 339–399

[45] S. Weis (2014) *Continuity of the maximum-entropy inference*, Communications in Mathematical Physics 330(3) 1263–1292

[46] S. Weis (2014) *The MaxEnt extension of a quantum Gibbs family, convex geometry and geodesics*, AIP Conf. Proc. 1641 173–180.

[47] S. Weis, A. Knauf (2012) *Entropy distance: New quantum phenomena*, J Math Phys 53(10) 102206

[48] E. H. Wichmann (1963) *Density matrices arising from incomplete measurements*, Journal of Mathematical Physics 4(7) 884–896

[49] D. Zhou (2008) *Irreducible multiparty correlations in quantum states without maximal rank*, Physical Review Letters 101(18) 180505

Stephan Weis
Departamento de Matemática
Instituto de Matemática, Estatística e Computação Científica
Universidade Estadual de Campinas
Campinas - SP - 13083-859
Brazil
e-mail: maths@stephan-weis.info